Journal of Sports Sciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjsp20

The use of 2D and 3D information in a perceptual-cognitive judgement task

Koen Put[a], Johan Wagemans[b], Jochim Spitz[a], Manuel Armenteros Gallardo[c], A. Mark Williams[d] & Werner F. Helsen[a]

[a] Department of Kinesiology, Movement Control and Neuroplasticity Research Group, KU Leuven, Belgium

[b] Laboratory of Experimental Psychology, KU Leuven, Belgium

[c] Department of Journalism and Audiovisual Communication, University Carlos III of Madrid, Spain

[d] Centre for Sports Medicine and Human Performance, Brunel University, United Kingdom
Published online: 23 May 2014.

Journal of Sports Sciences, 2014
Vol. 32, No. 18, 1688–1697, http://dx.doi.org/10.1080/02640414.2014.912760

The use of 2D and 3D information in a perceptual-cognitive judgement task

KOEN PUT[1], JOHAN WAGEMANS[2], JOCHIM SPITZ[1], MANUEL ARMENTEROS GALLARDO[3], A. MARK WILLIAMS[4] & WERNER F. HELSEN[1]

[1]*Department of Kinesiology, Movement Control and Neuroplasticity Research Group, KU Leuven, Belgium,* [2]*Laboratory of Experimental Psychology, KU Leuven, Belgium,* [3]*Department of Journalism and Audiovisual Communication, University Carlos III of Madrid, Spain and* [4]*Centre for Sports Medicine and Human Performance, Brunel University, United Kingdom*

(*Accepted 1 April 2014*)

Abstract

We examined whether the use of three-dimensional (3D) simulations in an off-field offside decision-making task is beneficial compared to the more widely available two-dimensional (2D) simulations. Thirty-three assistant referees, who were all involved in professional football, participated in the experiment. They assessed 40 offside situations in both 2D and 3D formats using a counterbalanced design. A distinction was made between offside situations near (i.e., 15 m) and far (i.e., 30 m) from the touchline. Subsequently, a frame recognition task was performed in which assistant referees were asked to indicate which of the five pictures represented the previous video scene. A higher response accuracy score was observed under 3D (80.0%) compared to 2D (75.0%) conditions, in particular for the situations near the touchline (3D: 81.8%; 2D: 72.7%). No differences were reported between 2D and 3D in the frame recognition task. Findings suggest that in highly dynamic and complex situations, the visual system can benefit from the availability of 3D information, especially for relatively fine, metric position judgements. In the memory task, in which a mental abstraction had to be made from a dynamic situation to a static snapshot, 3D stereo disparities do not add anything over and beyond 2D simulations. The specific task demands should be taken into account when considering the most appropriate format for testing and training.

Keywords: *stereoscopic vision, decision-making, recognition, perceptual training, assistant referees*

Introduction

One of the most robust findings in the literature on expertise is that expert musicians and chess players (Ericsson, Krampe, & Tesch-Römer, 1993) as well as elite performers in individual (Starkes, Deakin, Allard, Hodges, & Hayes, 1996) and team sports (Helsen, Starkes, & Hodges, 1998) devote around 10,000 hours or 10 years of deliberate practice to attain expert performance. As a result of this prolonged engagement in sport-specific practice (Williams & Ford, 2008) and play activities (Côté, Baker, & Abernethy, 2007; Williams, Ward, Bell-Walker, & Ford, 2012), experts develop superior perceptual-cognitive skills, such as pattern recognition and the ability to pick up information from the movements of others (for a review, see Williams, Ford, Eccles, & Ward, 2011).

The importance of this engagement in deliberate practice has also been shown to apply to officials, such as referees and assistant referees in association football (MacMahon, Helsen, Starkes, & Weston, 2007). Catteeuw, Helsen, Gilis, and Wagemans (2009) reported that the accumulated practice hours, the practice hours per week, the number of matches officiated, and years of officiating are the strongest predictors of performance in football referees. An important distinction between data from officials compared to athletes is that the latter group accumulated significant practice hours in sport-specific deliberate play, whereas for officials, the majority of practice time was spent in physical training rather than in practicing key perception and decision-making skills. In addition, given that the number of really difficult instances is too limited to learn from real games only, the objective of the current study was to explore the relative benefits of using two-dimensional (2D) and three-dimensional (3D) simulations to increase the amount of task-specific training for assistant referees.

One option is to provide assistant referees with additional off-field decision-making experience outside the real-life matches using video-based training

Correspondence: Werner F. Helsen, Department of Kinesiology, Movement Control and Neuroplasticity Research Group, KU Leuven, Belgium. E-mail: werner.helsen@faber.kuleuven.be

(e.g., see Larkin, Berry, Dawson, & Lay, 2011; Mascarenhas, Collins, Mortimer, & Morris, 2005; Pizzera & Raab, 2012; Schweizer, Plessner, Kahlert, & Brand, 2011). In recent years, a more scientific approach to the training and preparation of referees and assistant referees has been adopted. Sports scientists and international football-governing bodies (i.e., Union Européenne de Football Association (UEFA) and Fédération Internationale de Football Association (FIFA)) have focused attention on the development of role-specific perceptual-cognitive training programmes to improve the level of refereeing. For example, it was shown that a four-week offside decision-making training intervention with video simulations and computer animations resulted in higher response accuracy and a substantial decrease of flag errors (FE) (i.e., assistant referees signal for offside while the attacker is not in an offside position) in off-field offside decisions (Catteeuw, Gilis, Jaspers, Wagemans, & Helsen, 2010; Catteeuw, Gilis, Wagemans, & Helsen, 2010b).

Furthermore, these off-field training programmes induced a positive learning effect to better deal with the perceptual consequences of the flash-lag effect. This perceptual illusion, whereby the receiving attacker is perceived ahead of his actual position at the moment of the pass, can be considered as the most important factor in explaining the majority of incorrect offside decisions (Baldo, Ranvaud, & Morya, 2002; Catteeuw, Gilis, García-Aranda, et al., 2010; Gilis, Helsen, Catteeuw, Van Roie, & Wagemans, 2009).

Until recently, there was no clear indication of the potential benefits of video-based training programmes to improve the overall performance of assistant referees in real-match incidents. Put, Wagemans, Jaspers, and Helsen (2013) showed that web-based offside training, using video simulations and computer animations, resulted in a positive and direct transfer to on-field offside decisions. The structure and content of this training intervention corresponded to the perceptual difficulties of real-match situations and helped the assistant referees to mediate and enhance their on-field offside decision-making skills.

An important issue was that assistant referees perceived the fidelity, which refers to the degree of similarity between the training task and the real task (Hays & Singer, 1989), greater for video simulations than for computer animations (Catteeuw, Gilis, Jaspers, et al., 2010). A distinct difference in viewing perspective ("ego" perspective in video simulations, i.e., the perspective of the assistant referee and "top-view" perspective in computer animations) can explain the differences in physical (how real the simulations are), functional (how similar to the real situation the simulations function), and psychological (how real the simulations are experienced) fidelity between both formats.

The ultimate aim is to create simulations that can improve the perceptual-cognitive mechanisms underpinning expert perception and performance in (offside) decision-making. An important consideration is the extent to which the practice/test scenario replicates the constraints of the actual (*in situ*) performance setting. Therefore, the use of realistic and life-like video simulations (i.e., with high fidelity) should be encouraged, improving the ecological validity and the measurement sensitivity of the task (Dicks, Button, & Davids, 2010; Roca, Williams, & Ford, 2013). Some examples of such applications can be found in the field of human factors examining aviation (Russo et al., 2005) and military combat decision-making training (Ward et al., 2008; Williams, Ericsson, Ward, & Eccles, 2008).

To better represent the complex and ever-changing circumstances typically characterising a sport context, 3D information may be considered to offer additional fidelity when compared to the "traditional" 2D images of offside situations used in previous studies (Catteeuw et al., 2010b; Catteeuw, Gilis, Jaspers, et al., 2010; Put, Baldo, Cravo, Wagemans, & Helsen, 2013). At present, 3D displays are used extensively to accelerate the learning process of surgeons during the acquisition and fine-tuning of different surgical skills (e.g., Alaraj et al., 2013; Christopher, William, & Cohen-Gadol, 2013; Heath & Cohen-Gadol, 2012; Smith et al., 2012).

While the use of 3D is already well integrated into the medical community, the practical utility of such displays within the context of perceptual-cognitive training in sports is yet to be explored. Therefore, the aim of this study was to examine whether the use of 3D offside simulations in an off-field offside decision-making task can offer advantages compared to the use of more widely available 2D simulations. As 3D can facilitate the judgement of absolute and relative distances (for a review, see Lambooij, IJsselsteijn, Fortuin, & Heynderickx, 2009), considered as critical components in offside decision-making, we expected a higher response accuracy under 3D when compared to 2D viewing conditions. In particular for the offside situations occurring close to the touchline, that is, in the "near-field", we anticipated a more natural depth perception resulting in a better performance in contrast to 2D. We hypothesised that the powerful depth cue, referred to as "stereopsis" and provided by binocular disparity, will appear more effective for the near condition, because more details can be observed due to a wider viewing angle (Howard & Rogers, 2002). In the far condition, however, the angle of view is reduced and the differences between both images (overlapping in one video, but separated with the glasses) will become less prominent.

After assessing the offside situations, the assistant referees had to recall the exact position of the receiving attacker relative to the offside line at the moment of the pass. In similar vein, we predicted a significantly higher score on this frame recognition task in 3D in comparison with 2D.

Method

Participants

Altogether, 33 assistant referees (mean age = 38.4 years, s = 6.4) agreed to participate. All participants were working in professional football in Belgium. The sample size was deemed to provide sufficient power based on an *a priori* calculation (Cohen, 1992), using effect sizes from previous studies in offside decision-making (Catteeuw, Gilis, Jaspers, et al., 2010; Catteeuw, Helsen, Gilis, Van Roie, & Wagemans, 2009). On average, they had 8.7 years (s = 3.9) experience as an assistant referee. All participants were completely naïve with respect to the particular hypotheses being tested. The study took place as part of one of the weekly scheduled physical training sessions at the Faculty of Kinesiology and Rehabilitation Sciences in Leuven. Written consent was obtained from the Belgian Football Federation Referees' Committee and from each referee prior to testing according to the Declaration of Helsinki. In addition, the experimental protocol received approval from the local ethics committee at the KU Leuven.

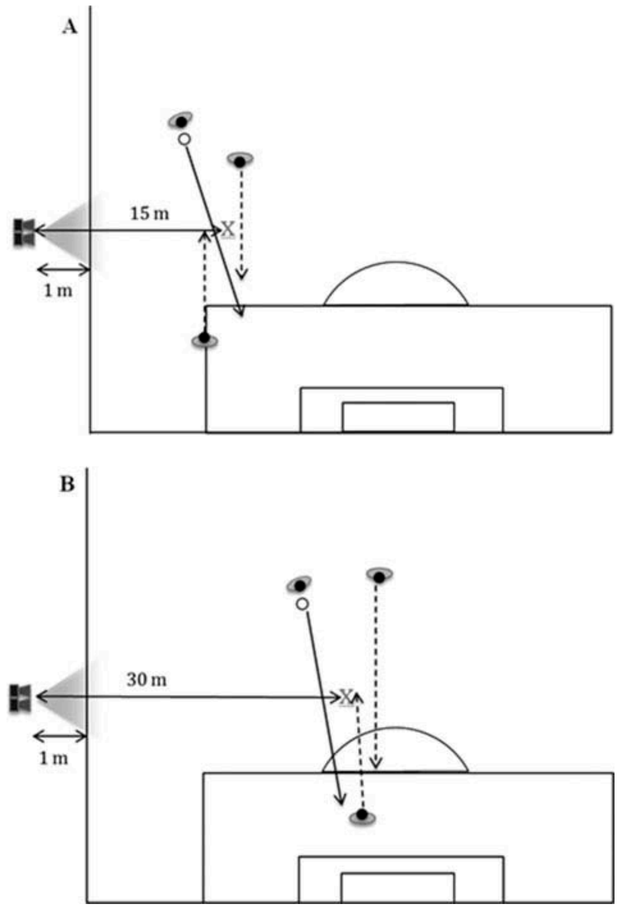

Figure 1. Representative examples of offside situations occurring at 15 m from the camera (A, "near") and 30 m from the camera (B, "far"). The movement patterns of the players and the ball direction are also indicated by the dashed and full lines, respectively.

Apparatus

A sample of youth elite players, aged between 16 and 18 years, simulated typical offside situations with two attackers and one defender. The players, all actively involved in competitive football, performed the situations as realistically as possible. Prior to the actual recording, the players received detailed instructions, and several training opportunities were provided to familiarise them with the task requirements. Next, the offside situations were recorded with two digital high-definition video cameras (DSLR Canon 550D, 35-mm focal length) at 60 frames per second with a 720p resolution. Both cameras were placed 1 m from the touchline, representing a match-like perspective. The distance between both cameras and the second-last defender was either 15 m, which was considered as "near" (Figure 1A), or 30 m, regarded as "far" (Figure 1B), relative to each other. The initial positions of the defender and the attackers were kept constant by placing cones on the field of play. In addition, the speed of the players was not measured directly, but the players carried out a trajectory (i.e., a certain distance between two cones) within a given time. By doing so, the players were able to maintain the same speed throughout the entire duration of the offside footage.

Out of a total of 600 offside actions that were recorded, 40 high-quality situations were chosen. First, only those situations in which the crossover between the attacker and the second-last defender took place exactly in front of the camera were selected. Second, the ball and the three players (two attackers and one defender) had to remain visible at all times. Importantly, this selection was made together with a former FIFA assistant referee who is now an expert assistant referee instructor. Consequently, the clips were divided into 4 categories of 10 situations according to the spatial positions of the attacker (Figure 2).

The ratio of onside versus offside situations was 75% versus 25%, respectively, across both viewing conditions. This rationale was in agreement with previous research and has been discussed with the technical instructors of the European football

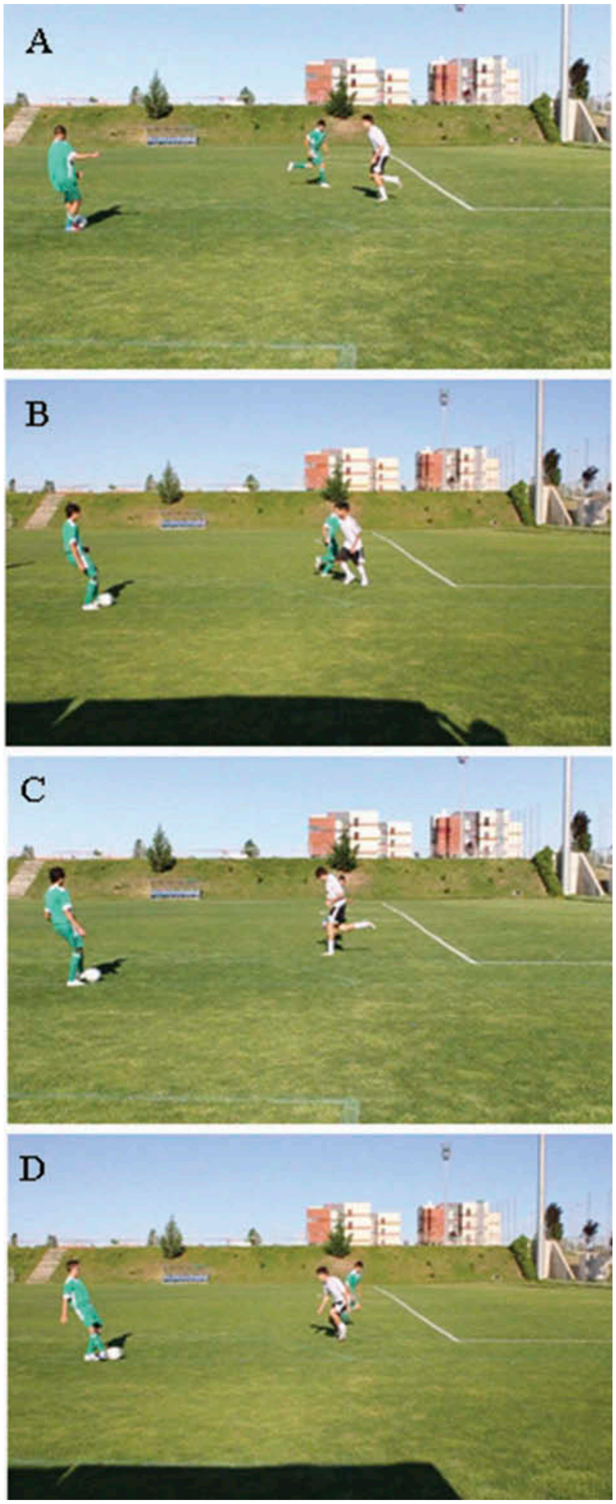

Figure 2. Illustration of the four spatial positions of the receiving attacker in relation to the second-last defender: (A) attacker clearly in front of the offside line: >1 m: onside; (B) attacker slightly in front of the offside line: 0–1 m: onside; (C) attacker on the offside line: 0 m: onside; (D) attacker slightly ahead of the offside line: 0–1 m: offside. © [University Carlos III of Madrid]. Reproduced by permission of Manuel Armenteros Gallardo.

governing body UEFA. From all of these situations, 22 offside situations were finally selected from the "near" condition and 18 from the "far" condition. The offside situations were digitised and edited using the software program Final Cut Pro 6.0.6 (Apple, Inc., USA).

Task and procedure

All assistant referees judged 80 offside situations (40 in 2D and 40 in 3D) represented in a video-simulation format from an in-game perspective to ensure that critical and relevant information could be identified adequately (Johnson, 2006; Mann, Farrow, Shuttleworth, & Hopwood, 2009). In line with previous research regarding offside decision-making (e.g., Put, Wagemans, et al., 2013), the laboratory offside assessment task consisted of two different parts. First, the assistant referees judged the offside situation as accurately as possible and indicated whether it was offside or not. The definition of an offside position states that "A player is in an offside position if he is nearer to his opponents' goal line than both the ball and the second-last opponent" (Law 11; FIFA, 2013). This part of the task had to be completed within a time window of 5 s after the final and decisive pass. Participants entered their answer on a test form. Second, the assistant referees were asked to mark on the same test form the correct frame that corresponds with the spatial position of the attacker and second-last defender at the exact moment the ball was passed (i.e., frame recognition) (Figure 3). The frame that perfectly matches with the moment of the pass had to be chosen within a time frame of 10 s.

Immediately after each test session, the assistant referees were asked to answer three questions by indicating a mark on a 10-cm line using a visual analogue scale. The questions included three opinions about both 2D and 3D video simulations: (1) How much confidence do you have in your answers? (0 = no confidence, 10 = full confidence); (2) How do you perceive the level of difficulty of the test? (0 = very easy, 10 = very difficult); and (3) How do you perceive the level of similarity (i.e., psychological fidelity) between the offside situations used in this experiment and the offside situations in a real match? (0 = not similar, 10 = similar).

After the assistant referees received standardised instructions regarding the procedure, participants were randomly assigned to one of the two groups; the first group (n = 17) started with the assessment of 40 offside situations in 2D format, whereas the second group (n = 16) started at the same time with 40 offside situations in 3D format. As such, the starting condition was counterbalanced across participants. The test sessions consisted of the same video clips, but displayed in a randomised order. In other words, each assistant referee assessed the same offside situation twice, both in 2D and in 3D. The

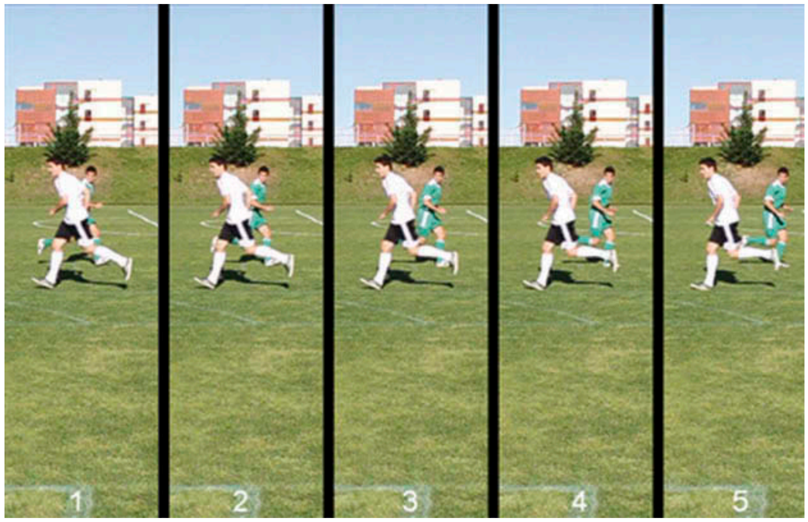

Figure 3. Screenshot of the five probe frames: the editing process started with seven positions for each offside situation with exactly a one-frame difference. Three frames were taken before the pass (–3, –2, –1): one (0) at the moment of the pass and three after the pass was given (+1, +2, +3). Eventually, the seven frames were reduced to a selection of three ranges of five frames (–3 to +1, –2 to +2, or –1 to +3). After each offside situation, one of these three ranges was randomly shown. © [University Carlos III of Madrid]. Reproduced by permission of Manuel Armenteros Gallardo.

set of 40 clips was divided into 4 series of 10 clips. A 2-min break was given after the first group of 10 clips, a 5-min break after 20 clips, and another 2-min break after 30 clips. All participants were requested not to discuss the offside clips amongst each other during the entire test session, including the various breaks.

With respect to the assessment of the 3D offside simulations, the assistant referees were equipped with red and cyan colour-coded anaglyph glasses. These glasses ensure that the two coloured 3D video simulations are transmitted as separate image streams to the visual cortex, where they are fused and perceived as a 3D scene. The assistant referees were given the opportunity to adapt to the 3D displays with a 4-min introductory animation prior to the actual offside simulations being shown. After each 10th offside situation, the assistant referees were allowed to remove their glasses to minimise and reduce the adverse effects of watching 3D with anaglyphs.

For both test sessions, the assistant referees were seated within the edges of the screen (3.20 × 4.20 m) to keep the viewing angle as small as possible. Also, the sitting position was kept identical relative to the screen across the sessions. Both test sessions took place in equally darkened rooms to guarantee optimal viewing settings.

Data analysis

Offside decision-making task

Response accuracy: The percentage of correct and incorrect decisions was calculated (40 trials = 40 points = 100% accuracy for both 2D and 3D). To further explore the response accuracy of the offside decision-making task, a distinction was made between situations that occurred near to the touchline (i.e., 15 m) and situations that took place far from the touchline (i.e., 30 m). When the assistant referees made an incorrect decision, a distinction was made between a FE and a non-flag error (NFE). The assistant referee indicated offside while the attacker was in an onside position (i.e., FE), or the referee decided no offside when the attacker was in an offside position (i.e., NFE).

Frame recognition task

Response accuracy: The percentage of correctly recognised frames was calculated for every offside situation (40 trials = 40 points = 100% accuracy for both 2D and 3D).

Weighted mean: This parameter was calculated by multiplying the proportion of responses at a given probe position by that probe's frame difference (–3, –2, –1, 0, +1, +2, or +3) from the correct probe (i.e., 0). These products are then added and divided by the total number of responses. If, for example, the weighted mean is >0, this means the assistant referees chose, in general, a frame following the correct answer (i.e., where the attacking player has moved further to the right than in the actual position at the time of the pass). If the weighted mean is <0, this means that in general a frame preceding the correct answer was chosen (i.e., where the attacking player has not moved to the right as much). For example, if an assistant referee chooses 5 times –3, 6 times –2, 8 times –1, 10 times 0 (correct judgement), 5 times +1, 4 times +2, and 2 times +3, then the weighted meanis calculated as follows: (–3 × 5) + (–2 × 6) + (–1 × 8) + 0 + 5 + 8 + 6 = –16. Expressed as a proportion, the weighted mean is –0.4 (–16/40) in this case.

Consistency: This parameter was calculated between the response accuracy on the video simulations on the one hand and the frame recognition task on the other hand. This was done to investigate whether the assistant referees chose an appropriate frame corresponding with their decision offside or no offside.

Statistical analysis

First, the results of the offside decision-making task violated the assumption of normality as tested by the Shapiro-Wilk normality test. Therefore, the Wilcoxon signed-rank test was used to compare results on the 2D and 3D the formats (expressed in median values). Bonferroni corrections for multiple comparisons were made. A *P*-value of <0.01 was considered significant. Effect sizes were calculated as $r = \frac{z}{\sqrt{N}}$. Effect sizes below the 0.30 criterion represent a small to medium effect, whereas an effect size above the 0.50 threshold indicates a large effect. Second, the frame recognition parameters (i.e., response accuracy, weighted mean, and consistency) were analysed using a 2 × 2 repeated measures ANOVA with format (2D and 3D) and distance (near and far) as within-participant variables. Significant main effects were further explored using Tukey *post hoc* procedures. Effect sizes were reported as partial-eta-square (η_p^2), only for *F*-values >1. Third, the parameters of the visual analogue scales (confidence, difficulty level, and psychological fidelity) were examined using dependent *t*-tests. Cohen's *d* was calculated as a measure of effect size (*d* = 0.2, small effect; *d* = 0.5, medium effect; *d* = 0.8, large effect). All statistical analyses were performed with STATISTICA 11.0 (StatSoft, Inc., USA). Unless otherwise stated, the *P*-value was set at 0.05.

Results

Offside decision-making task

Response accuracy: The results of a Wilcoxon signed-rank test showed that the overall response accuracy in 3D (Mdn = 80.0%, *s* = 4.76) was significantly higher compared to 2D (Mdn = 75.0%, *s* = 4.85) (*z* = 3.09, *T* = 74.5, pcorr <0.01, *r* = 0.54). The decision-making accuracy was analysed in further detail by distinguishing the offside situations as "near" (15 m from the touchline) and "far" (30 m from the touchline). No significant differences

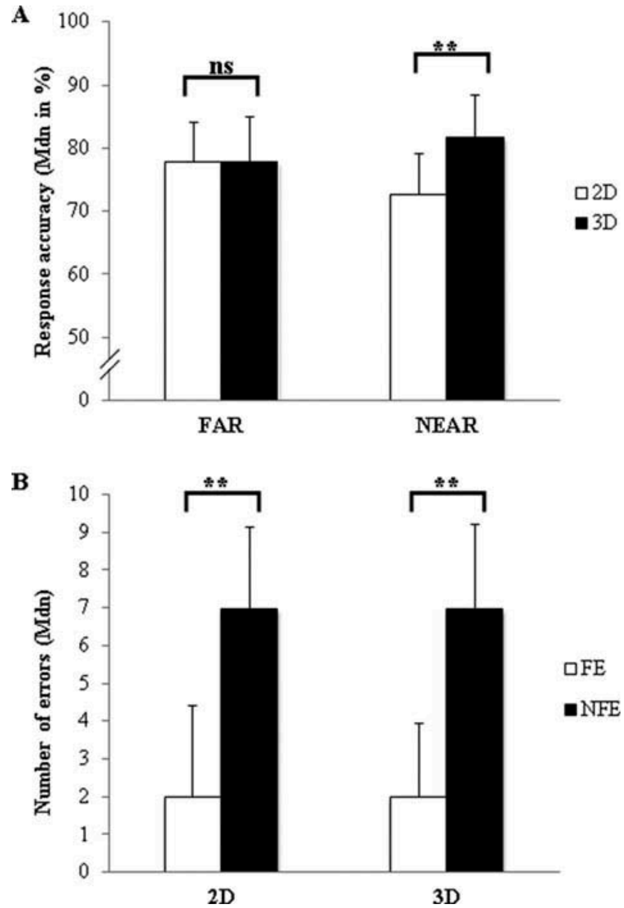

Figure 4. Response accuracy (A) and the type of incorrect decisions (flag errors (FE)/non-flag errors (NFE)) (B) during the offside decision-making task (**pcorr < 0.01).

between 3D (Mdn = 77.8%, *s* = 7.22) and 2D (Mdn = 77.8%, *s* = 6.22) were found for the far condition (*z* = 1.16, *T* = 119.5, *P* = 0.25, *r* = 0.20). For the near condition, however, response accuracy in 3D (Mdn = 81.8%, *s* = 6.49) was significantly higher compared to 2D (Mdn = 72.7%, *s* = 6.33) (*z* = 2.61, *T* = 96.5, pcorr <0.01, *r* = 0.45) (see Figure 4A).

As shown in Figure 4B, there were significantly more NFE than FE in both 2D (NFE: Mdn = 7.0, *s* = 2.13; FE: Mdn = 2.0, *s* = 2.41) (*z* = 4.07, *T* = 40.5, pcorr <0.01, *r* = 0.71) and 3D formats (NFE: Mdn = 7.0, *s* = 2.20; FE: Mdn = 2.0, *s* = 1.95) (*z* = 4.16, *T* = 47.5, pcorr <0.01, *r* = 0.72).

Frame recognition task

Table I presents the main results of the frame recognition task for both 2D and 3D formats.

Response accuracy: As can be seen from Table I, the 2 × 2 (format: 2D, 3D; distance: near, far) repeated measures ANOVA for response accuracy revealed no main effects of format [*F*(1,32) = 0.103, *P* = 0.75] and distance [*F*(1,32) = 0.018, *P* = 0.89]. The

Table I. Results of the frame recognition task (response accuracy, weighted mean, and consistency).

	2D	3D
Response accuracy (%)		
Overall	26.1 ± 7.02	25.8 ± 7.51
Far	26.1 ± 9.97	25.6 ± 10.48
Near	26.2 ± 8.28	25.9 ± 9.05
Weighted mean		
Overall	−0.24 ± 0.37	−0.22 ± 0.46
Far	−0.10 ± 0.46*	−0.10 ± 0.58*
Near	−0.35 ± 0.38*	−0.33 ± 0.44*
Consistency		
Overall	0.87 ± 0.05	0.89 ± 0.05
Far	0.91 ± 0.06*	0.93 ± 0.07*
Near	0.84 ± 0.06*	0.85 ± 0.06*

Note: *pcorr < 0.01.

format × distance interaction effect [$F(1,32) = 0.006$, $P = 0.95$] was also not significant. In contrast to the response accuracy on the offside decision-making task, no significant differences were found on the frame recognition task between 2D and 3D, regardless of the distance between the touchline and the second-last defender.

Weighted mean: When investigating the weighted mean, there was no main effect for format [$F(1,32) = 0.096$, $P = 0.76$] and no interaction between format × distance [$F(1,32) = 0.042$, $P = 0.83$], but there was a main effect for distance [$F(1,32) = 17.751$, $P < 0.001$, $\eta_p^2 = 0.36$]. Frames preceding the correct answer were significantly ($P < 0.05$) more likely to be chosen than the frames following the correct answer in both the 2D (preceding: mean value 17.6, s = 4.90; following: mean value 11.9, s = 4.33) and the 3D formats (preceding: 17.3 ± 5.92; following: mean value 12.4, s = 5.84), highlighting the overall bias towards NFE. Table I shows that the values of the weighted mean for the near condition are even more negative compared to the far condition in both 2D and 3D (pcorr < 0.01).

Consistency: In analogy with the results of the weighted mean, the analysis of consistency showed no main effect for format [$F(1,32) = 2.34$, $P = 0.14$, $\eta_p^2 = 0.068$] or for the interaction between format × distance [$F(1,32) = 0.034$, $P = 0.85$], but there was a main effect for distance [$F(1,32) = 38.29$, $P < 0.001$, $\eta_p^2 = 0.545$]. Overall, the consistency is higher for the far condition than for the near condition in both the 2D and the 3D formats (pcorr < 0.01).

Visual analogue scales

The results of various dependent t-tests showed no significant differences between the 2D and the 3D formats regarding confidence (2D: mean value 74.45, s = 11.53; 3D: mean value 77.24, s = 11.14, $t = -1.04$, $P = 0.30$, $d = 0.25$) and perceived level of difficulty (2D: mean value 75.03, s = 16.99; 3D: mean value 71.03, s = 12.26, $t = 1.42$, $P = 0.16$, $d = 0.27$). However, there was a significant difference for psychological fidelity (2D: mean value 62.94, s = 17.73; 3D: mean value 70.0, s = 19.66, $t = -2.32$, $P < 0.05$, $d = 0.37$); participants experienced the 3D more similar to the real-match incidents when compared to 2D.

Discussion

To understand better the perceptual-cognitive constructs underlying expert perception and performance, it is necessary to develop experimental tasks in which the specificity and complexity of real-life situations are closely reproduced (Ericsson & Williams, 2007). In other words, task paradigms with high fidelity are preferred over simplistic and manufactured tasks with a stronger emphasis on internal rather than external validity (Dhami, Hertwig, & Hoffrage, 2004; Dicks, Davids, & Button, 2009; Roca et al., 2013). To provide more realistic and reproducible domain-specific situations, we employed 3D animated offside scenarios. To the best of our knowledge, no researchers to date have yet examined the added value of such displays in an off-field sports decision-making task.

In line with the main hypothesis, the results of this study showed that the overall response accuracy for the offside decision-making task was significantly higher under 3D compared with 2D viewing conditions. In highly dynamic and complex situations with many relevant and simultaneously moving objects, the visual system can benefit from having access to additional information available in 3D compared with 2D, in particular for relatively near position judgements (for a review, see St. John, Cowen, Smallman, & Oonk, 2001). This finding supports previous reports that suggest that the use of 3D displays benefits the human information processing system, for example, in the field of air traffic control, military personnel, and surgery (Ward, Williams, & Hancock, 2006). To analyse response accuracy in more detail, a distinction was made between offside situations near (i.e., 15 m) and far (i.e., 30 m) from the touchline. As expected, the results showed that the assistant referees performed better in the near condition in 3D, since the viewing angle is wider and the differences between both images become more prominent. In the far condition, however, retinal disparities are smaller and the 3D advantage is (partially) eliminated.

Surprisingly, no differences were observed between 2D and 3D for the response accuracy on the frame recognition task. In line with previous studies (e.g., Catteeuw et al., 2010b; Catteeuw, Gilis, Jaspers, et al., 2010), the assistant referees scored relatively low on this task, in both the 2D and the 3D formats, indicating a high difficulty level. When recalling the spatial positions of the attacker and the second-last defender (cf. "mental picture" of the offside situation), 3D stereo disparities did not seem to have any added value. It might be suggested that the perception of different movement and speed patterns (cf. offside situation) required another way of processing information compared to the perception of still images (cf. frame recognition) (Smallmann, St. John, Oonk, & Cowen, 2001; Tory, Möller, Atkins, & Kirkpatrick, 2004). In this specific recognition task, assistant referees are asked to select a static snapshot to represent the correct positions of the objects at the moment of the final pass. By doing so, a mental abstraction has to be made from a dynamic situation to a static snapshot. In this step away from a complicated spatiotemporal event in space to a distance judgement in a snapshot, the original 3D perception during the encoding stage is probably not helpful in the recognition stage. Although in the 3D condition the assistant referees are still using their 3D goggles in the static image recognition task, this appears not to have any added value.

It is clear that our study also addresses the issue of which display technique fits within the specific tasks demands (i.e., offside decision-making vs. frame recognition). A number of researchers have found benefits for 3D displays compared to 2D, whereas others have reported mixed or absent benefits of 3D (for an overview, see St. John et al., 2001). The results seemed to depend more on the specific requirements of each task paradigm and test set-ups than on the nature of the display formats themselves. Therefore, the task demands should be seriously taken into account when considering the most appropriate format for testing and training.

A substantial amount of literature has already been published focusing on the mechanisms underlying decision-making in offside situations, such as shift of gaze (Belda Maruenda, 2004), optical error (Oudejans et al., 2000, 2005), and perceptual error (Baldo et al., 2002; Catteeuw et al., 2010b; Catteeuw, Gilis, Jaspers, et al., 2010; Gilis et al., 2009; Put, Wagemans, et al., 2013). In this study, an overall bias existed towards NFE compared to FE in both 2D and 3D viewing conditions. This finding is in line with Catteeuw, Gilis, Wagemans, and Helsen (2010a), who only stated this result based on a video-based match analysis (i.e., 2D format). As such, this observation might support the idea of cognitive overcompensation. It has been demonstrated by Catteeuw et al. (2010b) that even a limited number of perceptual-cognitive training sessions can help assistant referees to develop a cognitive compensation strategy to better deal with the illusive effect of flash-lag. It is generally known that expertise levels critically depend on the ability to overcome such perceptual-cognitive difficulties. However, the results from the present study reveal that assistant referees do not seem to make the required adjustment at all times. This strategy of cognitive compensation may have led to an overcorrection/overcompensation of their (incorrect) perception, resulting in a clear bias towards NFE (Put, Baldo, et al., 2013). As discussed by Catteeuw et al. (2010a), this result can also partially be explained by the specific instructions given by the international football governing bodies, such as FIFA and UEFA, to indicate no offside in case of doubt, increasing the possibility of making more NFE.

In line with the overall bias towards NFE, the negative values of the weighted mean, in both the 2D and the 3D formats, indicated that the assistant referees mainly chose the frames preceding the correct answer, as compared to the frames following the correct answer. This finding can support the aforementioned strategy of overcompensation, suggesting that assistant referees cognitively overcorrect their perceived recall of the "freeze" of the spatial positions of the players.

The results of the visual analogue scale showed significant differences between the 2D and the 3D formats for psychological fidelity. The assistant referees experienced the degree of similarity between the off-field task and the real task on the field of play higher in 3D when compared to 2D. Thus, in line with a major aim of the present study, this result clearly confirms the ecological validity of our task paradigm. With respect to confidence and perceived level of difficulty, one can critically argue that both can fluctuate from one situation to another. It can also be stated that self-report of confidence, particularly in the population of officials, can give an overestimation since they must deliver decisions confidently and are thus prone to report high confidence scores.

Further research is required to keep abreast of significant changes in technology leading to further enhancements in the quality of 3D. First, nowadays, the film industry is already working with twice the speed of the frames used in this study (120 frames per second instead of 60) and Ultra HD resolution (3840 pixels instead of 720) to obtain a more realistic perception of the ongoing situations. Second, since 3D situations clearly increase the fidelity of the offside clips, it would be interesting to investigate the additional training and transfer benefits to the real world of offside decision-making. Third, it is important to limit the "simulator-sickness-type" symptoms, such as eyestrain, nausea, and measurable changes in visual functions (Pölönen, Järvenpää,

& Bilcu, 2013), as reported by a part of the assistant referees after the 3D test sessions.

In conclusion, in this study, we revealed that 3D displays have an added value when compared to 2D displays with respect to the perception and decision-making performance of assistant referees. The 3D displays improved a variety of perceptual-cognitive skills in sports officiating, although these effects were task dependent. Findings can be of interest to a broad sport audience interested in decision-making under dynamic and temporally constrained environments. Finally, the present results contribute to the existing knowledge base regarding 2D and 3D simulations and their practical utility in examining the role of specific task demands.

References

Alaraj, A., Charbel, F. T., Birk, D., Tobin, M., Luciano, C., Banerjee, P. P., & Roitberg, B. (2013). Role of cranial and spinal virtual and augmented reality simulation using immersive touch modules in neurosurgical training. *Neurosurgery*, *72*, A115–A123. doi:10.1227/NEU.0b013e3182753093

Baldo, M. V. C., Ranvaud, R. D., & Morya, E. (2002). Flag errors in soccer games: The flash-lag effect brought to real life. *Perception*, *31*, 1205–1210. doi:10.1068/p3422

Belda Maruenda, F. (2004). Can the human eye detect an offside position during a football match? *British Medical Journal*, *329*, 1470–1472. doi:10.1136/bmj.329.7480.1470

Catteeuw, P., Gilis, B., García-Aranda, J.-M., Tresaco, F., Wagemans, J., & Helsen, W. (2010). Offside decision making in the 2002 and 2006 FIFA World Cups. *Journal of Sports Sciences*, *28*, 1027–1032. doi:10.1080/02640414.2010.491084

Catteeuw, P., Gilis, B., Jaspers, A., Wagemans, J., & Helsen, W. (2010). Training of perceptual-cognitive skills in offside decision making. *Journal of Sport & Exercise Psychology*, *32*, 845–861.

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010a). Offside decision making of assistant referees in the English Premier League: Impact of physical and perceptual-cognitive factors on match performance. *Journal of Sports Sciences*, *28*, 471–481. doi:10.1080/02640410903518184

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010b). Perceptual-cognitive skills in offside decision making: Expertise and training effects. *Journal of Sport & Exercise Psychology*, *32*, 828–844.

Catteeuw, P., Helsen, W., Gilis, B., Van Roie, E., & Wagemans, J. (2009). Visual scan patterns and decision-making skills of expert assistant referees in offside situations. *Journal of Sport & Exercise Psychology*, *31*, 786–797.

Catteeuw, P., Helsen, W., Gilis, B., & Wagemans, J. (2009). Decision-making skills, role specificity and deliberate practice in association football refereeing. *Journal of Sports Sciences*, 27, 1125–1136. doi:10.1080/02640410903079179

Christopher, L. A., William, A., & Cohen-Gadol, A. A. (2013). Future directions in 3-dimensional imaging and neurosurgery: Stereoscopy and autostereoscopy. *Neurosurgery*, *72*, A131–138. doi:10.1227/NEU.0b013e318270d9c0

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*, 98–101. 10.1111/1467-8721.ep10768783

Côté, J., Baker, J., & Abernethy, B. (2007). Practice and play in the development of sport expertise. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 184–202). New York, NY: Wiley.

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*, 959–988. doi:10.1037/0033-2909.130.6.959

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, & Psychophysics*, *72*, 706–720. 10.3758/APP.72.3.706

Dicks, M., Davids, K., & Button, C. (2009). Representative task designs for the study of perception and action in sport. *International Journal of Sport Psychology*, *40*, 506–524.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406. doi:10.1037/0033-295X.100.3.363

Ericsson, K. A., & Williams, A. M. (2007). Capturing naturally occurring superior performance in the laboratory: Translational research on expert performance. *Journal of Experimental Psychology: Applied*, *13*, 115–123. doi:10.1037/1076-898X.13.3.115

Fédération Internationale de Football Association (2013). *Laws of the game*. Zürich: FIFA.

Gilis, B., Helsen, W., Catteeuw, P., Van Roie, E., & Wagemans, J. (2009). Interpretation and application of the offside law by expert assistant referees: Perception of spatial positions in complex dynamic events on and off the field. *Journal of Sports Sciences*, *27*, 551–563. doi:10.1080/02640410802702178

Hays, R. T., & Singer, M. J. (1989). *Simulation fidelity in training system design: Bridging the gap between reality and training*. New York, NY: Springer.

Heath, M. D., & Cohen-Gadol, A. A. (2012). Intraoperative stereoscopic 3D video imaging: Pushing the boundaries of surgical visualisation and applications for neurosurgical education. *British Journal of Neurosurgery*, *26*, 662–667. doi:10.3109/02688697.2012.672057

Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practise. *Journal of Sport and Exercise Psychology*, *20*, 12–34.

Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth*. Toronto, ON: I. Porteous.

Johnson, J. G. (2006). Cognitive modeling of decision making in sports. *Psychology of Sport and Exercise*, *7*, 631–652. 10.1016/j.psychsport.2006.03.009

Lambooij, M., IJsselsteijn, W., Fortuin, M., & Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, *53*, 30201–30214. doi:10.2352/J.ImagingSci.Technol.2009.53.3.030201

Larkin, P., Berry, J., Dawson, B., & Lay, B. (2011). Perceptual and decision-making skills of Australian football umpires. *International Journal of Performance Analysis in Sport*, *11*, 427–437.

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, *25*, 65–78. doi:10.1080/02640410600718640

Mann, D. L., Farrow, D., Shuttleworth, R., & Hopwood, M. (2009). The influence of viewing perspective on decision-making and visual search behaviour in an invasive sport. *International Journal of Sport Psychology*, *40*, 546–564.

Mascarenhas, D. R. D., Collins, D., Mortimer, P., & Morris, R. L. (2005). A naturalistic approach to training accurate and coherent decision making in rugby union referees. *The Sport Psychologist*, *19*, 131–147.

Oudejans, R. R., Bakker, F. C., Verheijen, R., Gerrits, J. C., Steinbruckner, M., & Beek, P. J. (2005). How position and motion of expert assistant referees in soccer relate to the quality of their offside judgements during actual match play. *International Journal of Sport Psychology*, *36*, 3–21.

Oudejans, R. R., Verheijen, R., Bakker, F. C., Gerrits, J. C., Steinbrückner, M., & Beek, P. J. (2000). Errors in judging "offside" in football. *Nature*, *404*. doi:10.1038/35003639

Pizzera, A., & Raab, M. (2012). Does motor or visual experience enhance the detection of deceptive movements in football? *International Journal of Sports Science & Coaching*, *7*, 269–284. doi:10.1260/1747-9541.7.2.269

Pölönen, M., Järvenpää, T., & Bilcu, B. (2013). Stereoscopic 3D entertainment and its effect on viewing comfort: Comparison of children and adults. *Applied Ergonomics*, *44*, 151–160. doi:10.1016/j.apergo.2012.06.006

Put, K., Baldo, M. V. C., Cravo, A., Wagemans, J., & Helsen, W. (2013). Experts in offside decision making learn to compensate for their illusory perceptions. *Journal of Sport and Exercise Psychology*, *35*, 576–584.

Put, K., Wagemans, J., Jaspers, A., & Helsen, W. (2013). Web-based training improves on-field offside decision-making performance. *Psychology of Sport and Exercise*, *14*, 577–585. doi:10.1016/j.psychsport.2013.03.005

Roca, A., Williams, A. M., & Ford, P. R. (2013). Capturing and testing perceptual-cognitive expertise: A comparison of stationary and movement response methods. *Behavior Research Methods*, *46*, 173–177. doi:10.3758/s13428-013-0359-5

Russo, M. B., Kendall, A. P., Johnson, D. E., Sing, H. C., Thorne, D. R., Escolas, S. M., … Redmond, D. P. (2005). Visual perception, psychomotor performance, and complex motor performance during an overnight air refueling simulated flight. *Aviation, Space, and Environmental Medicine*, *76*, C92–C103.

Schweizer, G., Plessner, H., Kahlert, D., & Brand, R. (2011). A video-based training method for improving soccer referees' intuitive decision-making skills. *Journal of Applied Sport Psychology*, *23*, 429–442. doi:10.1080/10413200.2011.555346

Smallmann, H. S., St. John, M., Oonk, H. M., & Cowen, M. B. (2001). Information availability in 2D and 3D displays. *IEEE Computer Graphics and Applications*, *21*, 51–57. doi:10.1109/38.946631

Smith, R., Day, A., Rockall, T., Ballard, K., Bailey, M., & Jourdan, I. (2012). Advanced stereoscopic projection technology significantly improves novice performance of minimally invasive surgical skills. *Surgical Endoscopy*, *26*, 1522–1527. doi:10.1007/s00464-011-2080-8

Starkes, J. L., Deakin, J. M., Allard, F., Hodges, N. J., & Hayes, A. (1996). Deliberate practise in sports: What is it anyway? In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 81–106). Hillsdale, MI: Lawrence Erlbaum Associates.

St. John, M., Cowen, M. B., Smallman, H. S., & Oonk, H. M. (2001). The use of 2D and 3D displays for shape-understanding versus relative-position tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*, 79–98. doi:10.1518/001872001775992534

Tory, M., Möller, T., Atkins, M. S., & Kirkpatrick, A. E. (2004). Combining 2D and 3D views for orientation and relative position tasks. *Proceedings of the SIGCHI conference on human factors in computing systems*, 73–80. doi:10.1145/985692.985702

Ward, P., Farrow, D., Harris, K. R., Williams, A. M., Eccles, D. W., & Ericsson, K. A. (2008). Training perceptual-cognitive skills: Can sport psychology research inform military decision training? *Military Psychology*, *20*, SS71–SS102. doi:10.1080/08995600701804814

Ward, P., Williams, A. M., & Hancock, P. A. (2006). Simulation for performance and training. In K. A. Ercisson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 243–262). New York, NY: Cambridge University Press.

Williams, A. M., Ericsson, K. A., Ward, P., & Eccles, D. W. (2008). Research on expertise in sport: Implications for the military. *Military Psychology*, *20*, SS123–SS145. doi:10.1080/08995600701804863

Williams, A. M., & Ford, P. R. (2008). Expertise and expert performance in sport. *International Review of Sport and Exercise Psychology*, *1*, 4–18. doi:10.1080/17509840701836867

Williams, A. M., Ford, P. R., Eccles, D. W., & Ward, P. (2011). Perceptual-cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Applied Cognitive Psychology*, *25*, 432–442. doi:10.1002/acp.1710

Williams, A. M., Ward, P., Bell-Walker, J., & Ford, P. R. (2012). Perceptual-cognitive expertise, practice history profiles and recall performance in soccer. *British Journal of Psychology*, *103*, 393–411. doi:10.1111/j.2044-8295.2011.02081.x